Multiple extraction and voice in Toba Batak

Michael Yoshitaka Erlewine

National University of Singapore

`mitcho@nus.edu.sg`

Toba Batak has a Malay/Indonesian-type voice system and is thought to only allow extraction of one DP at a time (Cole and Hermon, 2008).

1. **Multiple, simultaneous extractions** to the left-periphery—including extraction of multiple DPs—is possible, under certain circumstances.
   - When multiple DPs are fronted, voice morphology tracks the DP moved to **immediately preverbal** position.

2. The pattern of possible multiple extractions motivates **a head-splitting view of the C-T connection** (Martinović, 2015; Aldridge, 2015): CT starts as a single head, but sometimes splits.
   - Different probes associated with C and T, but they **first probe together**.

- Often simply *Hata Batak* 'Batak language'
- Spoken in northern Sumatra, around Lake Toba
- Two million speakers, according to Ethnologue
- Data here from elicitation with two speakers in Singapore

U. Michigan Museum of Anthropology

Roadmap

Roadmap

Toba Batak exhibits a two-way voice alternation, similar to Malay/Indonesian languages: (PN = proper name marker)

(1) Schachter (1984a, p. 123):

    a. **Mang**-ida si Ria **si Torus**.
        ACT-see   PN Ria PN Torus

    b. **Di**-ida   si Torus **si Ria**.
        PASS-see PN Torus PN Ria

        'Torus saw Ria.'

☞ The **voice prefix** tracks the choice of **pivot** argument (here sentence-final). I refer to *maN-* (16a) as ACTIVE and *di-* (16b) as PASSIVE.

Voice in Toba Batak

Verb-initial order is the canonical declarative order, but more than one third of declaratives in some texts have a fronted pivot (Cumming, 1984):

(2) a. **Si Torus** [**mang**-ida si Ria _____].
         PN Torus ACT-see   PN Ria

     b. **Si Ria** [**di**-ida si Torus _____].
         PN Ria PASS-see PN Torus

         'Torus saw Ria.'

Cumming (1984) describes this fronting as associated with topichood and reports that such fronted topics are "overwhelmingly definite" or generic.

In transitive clauses, the DP argument that is *not* the pivot (Schachter's "internal noun phrase") must be strictly verb-adjacent:

(3) **Adding *nantoari* 'yesterday' to (16a,b):** (Schachter, 1984a, p. 125)

    a. (✓Nantoari) mang-ida (⋆) si Ria (✓) si Torus (✓).
                ACT-see   PN Ria   PN Torus

    b. (✓Nantoari) di-ida   (⋆) si Torus (✓) si Ria (✓).
                PASS-see   PN Torus   PN Ria

      'Torus saw Ria yesterday.'

Emmorey (1984) shows that this argument always forms a unit together with the verb for the purposes of nuclear stress assignment.

Voice in Toba Batak: Extracting a DP

If a DP is fronted, it must be the pivot:

(4) **Actor *wh*-question:**

    a. ✓ ***Ise*** [**mang**-allang pinahan-on _____]?
          who ACT-eat   pork-this

    b. * *Ise* [**di**-allang _____ **pinahan-on**]?
          who PASS-eat   pork-this

        'Who ate this pork?'

(5) **Patient *wh*-question:**

    a. * *Aha* [**ma**-nuhor _____ **si Poltak**]?
          who ACT-buy   PN Poltak

    b. ✓ ***Aha*** [**di**-tuhor si Poltak _____]?
          who PASS-buy PN Poltak

        'What did Poltak buy?'

Voice in Toba Batak: Extracting a non-DP

Fronting of non-DPs does not interact with voice; both voices are possible, with corresponding postverbal word order:

(6) **Oblique *wh*-question:**

    a. ✓ [*Tu ise*] [**ma**-nuhor buku **si Poltak**]?
         DAT who ACT-buy   book PN Poltak

    b. ✓ [*Tu ise*] [**di**-tuhor si Poltak **buku**]?
         DAT who PASS-buy   PN Poltak book

       '[For who] did Poltak buy the book?'

(4–6) are my examples but Clark (1984, 1985) and Cole and Hermon (2008) describe the same pattern.

Cole and Hermon (2008)

Based on such facts, Cole and Hermon (2008) argue for a **V(oice)P-fronting** analysis for Toba Batak clauses:

- The non-pivot DP argument, if there is one, stays in-situ in VoiceP;
- All other arguments are moved out of VoiceP;
- VoiceP remnant-moves, freezes;
- ⇒ The non-pivot DP argument will be adjacent to the verb and cannot subsequently move

Related to more general questions about the derivation of verb-initiality; see also discussion in Chung (2008).

$\overline{\text{A}}$-movements: *wh*-DPs

Two types of $\overline{\text{A}}$-movements will be relevant here: **wh-movement** and **focus movement**.

Wh-words prefer to front, but can stay in-situ. *Wh*-in-situ is not an echo question, as diagnosed by question embedding:

(7) **True optional *wh*-movement:**

    a. Hu-boto   [*ise* [mang-allang pinahan]].
        PASS.1sg-know who ACT-eat   pork

    b. Hu-boto   [mang-allang pinahan *ise*].
        PASS.1sg-know ACT-eat   pork   who

    c. Hu-boto   [di-allang *ise*   pinahan].
        PASS.1sg-know PASS-eat   who pork
        'I know [who ate the pork].'

(8)   ***Wh*-movement is optional for adjuncts too:**

   a. *Andigan* ma-nuhor buku si Poltak?
       when   ACT-buy   book PN Poltak

   b. Ma-nuhor buku si Poltak *andigan*?
       ACT-buy   book PN Poltak when

   c. Ma-nuhor buku *andigan* si Poltak?
       ACT-buy   book when   PN Poltak

       'When did Poltak buy the book?'

(Passive variants all possible, with positions of Poltak and book reversed.)

Only-phrases are also best when fronted:

(9) **Focus-fronting preferred but both ok:**

    a. [Holan si Poltak] [mang-allang indahan _____].
       only   PN Poltak ACT-eat   rice

    b. Mang-allang indahan [holan si Poltak].
       ACT-eat   rice   only   PN Poltak

       'Only POLTAK ate rice.'

Roadmap

Multiple extractions: *wh*-DP + DP

Q1: Can you front two DPs at the same time?

A1: At first glance, no.

(10) ***Wh*-actor, regular DP patient:**
'Who ate the pork?'

    a.   *Ise* [mang-alang pinahan _____]?
          who ACT-eat   pork

    b.   Pinahan-on [di-allang *ise*   _____]?
          pork-this   PASS-eat who

    c. * *Ise* pinahan-on [mang/di-allang ___ ___]?
        who pork-this   ACT/PASS-eat

Cole and Hermon (2008, p. 183) discuss data such as (10c, 11c) and say this is predicted by their account.

Q1: Can you front two DPs at the same time?

A1: At first glance, no.

(11)   ***Wh*-patient, regular DP actor:**
      'What did Poltak buy?'

     a.   *Aha*   [di-tuhor si Poltak _____]?
            what PASS-buy   PN Poltak

     b.   Si Poltak [ma-nuhor *aha*   _____]?
            PN Poltak ACT-buy   what

     c.   * *Aha*   si Poltak [maN/di-tuhor __ __]?
            what PN Poltak ACT/PASS-buy

Cole and Hermon (2008, p. 183) discuss data such as (10c, 11c) and say this is predicted by their account.

Q2: But what if they're both $\overline{\text{A}}$-operators that prefer to front?

A2: They can both be fronted!

(12) **_Wh_-actor, _only_ patient:**
'Who ate only rice/pork?'

    a. *Ise* [mang-allang holan indahan _____]?
       who ACT-eat   only rice

    b. Holan pinahan [di-allang *ise* _____]?
       only pork   PASS-eat who

    c. *Ise* holan pinahan [{*mang/✓di}-allang __ __]?
       who only pork   {*ACT/✓PASS}-eat

Q2: But what if they're both A̅-operators that prefer to front?

A2: They can both be fronted!

(13)   ***Wh*-patient, *only* actor:**

'What did only Poltak eat?'

   a. *Aha*   [di-allang holan si Poltak _____]?
     what PASS-eat   only   PN Poltak

   b. Holan si Poltak [mang-allang *aha*   _____]?
     only   PN Poltak ACT-eat   what

   c. *Aha*   holan si Poltak [{✓mang/*di}-allang ___ ___]?
     what only   PN Poltak {✓ACT/*PASS}-eat

Q3: What about non-DP *wh*s? I remember those don't interact with voice.

A3: I'm glad you asked!

(14) **Non-DP *wh*, regular DP:**

    a.   *Andigan* buku-i   [{*maN/✓di}-tuhor ho __ __]?
          when   book-that {*ACT/✓PASS}-buy   2sg
          'When did you buy that book?'

    b.   *Andigan* si Poltak [{✓maN/*di}-tuhor buku __ __]?
          when   PN Poltak {✓ACT/*PASS}-buy   book
          'When did Poltak buy the book?'

Summary

(15) **Summary:**

    a.   * DP[*wh*] DP V…   (10–11)

    b.   ✓ DP[*wh*] **DP**[*only*] V…   (12–13)

    c.   ✓ Non-DP[*wh*] **DP** V…   (14)

Lesson 1: The non-pivot DP (internal noun phrase) *can* be moved, in certain circumstances, contra Cole and Hermon (2008).

Lesson 2: Voice tracks the choice of *immediately preverbal* DP.

Roadmap

Multiple extractions and voice

Recall that when multiple DPs are extracted, **voice tracks the *immediately preverbal* DP.**

$\Rightarrow$ The pivot DP is fronted first.

☞ The pivot DP is in a designated position (Guilfoyle, Hung, and Travis, 1992, a.o.) at the edge of the lower phase. DP probing from above will find the pivot first.

Voice

(16) **Working assumptions for voice (Erlewine, Levin, and Van Urk, 2015, to appear, in progress):**

a. One DP (the pivot) is attracted to a designated position (but may be pronounced low or to the right)

b. Voice morphology tracks this choice of pivot.

c. DPs need licensing (abstract Case):

- the pivot DP must be licensed from above (nominative)
- one DP (the non-pivot) can be licensed by PF adjacency with the verb (Levin, 2015, and references there)
- ⇒ this is the source of strict verb-adjacency for the non-pivot argument (when postverbal)
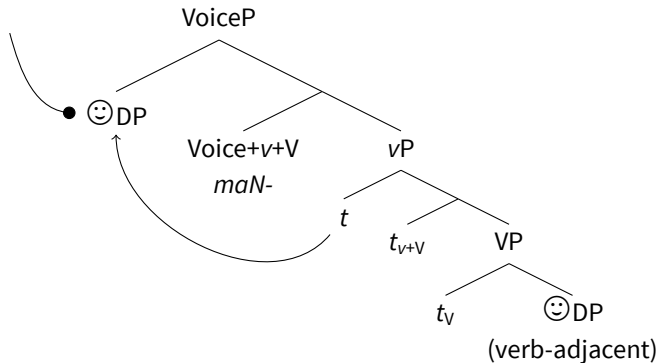
The voice details in (16) could conceivably be swapped out for different approaches to voice morphology.

Voice in Batak

VoiceP is the lower phase; actors are generated in Spec,*v*P below Voice (pace Legate, 2014). **The pivot is Spec,VoiceP** (pronounced to the right).
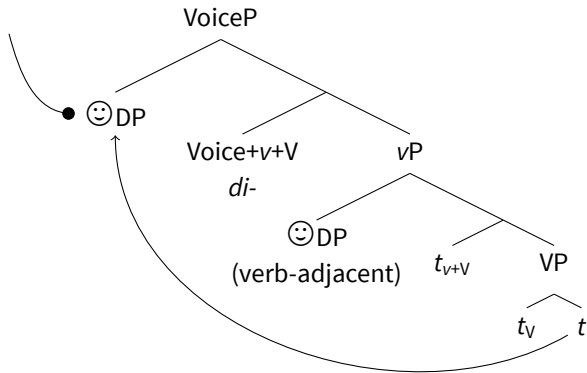
<u>Active voice:</u>

VoiceP
- ☺DP
- Voice+*v*+V *maN-*
- *v*P
  - *t*
  - *t*$_{v+V}$
  - VP
    - *t*$_V$
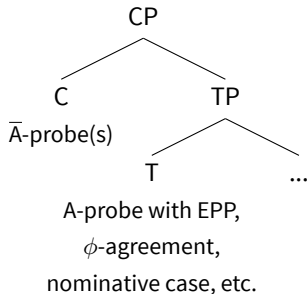   - ☺DP (verb-adjacent)

Voice in Batak

VoiceP is the lower phase; actors are generated in Spec,vP below Voice (pace Legate, 2014). **The pivot is Spec,VoiceP** (pronounced to the right).

Passive voice:

(17) **Traditional division of labor:** (Chomsky, 1986, a.o.)

   a. C: $\overline{\text{A}}$-movement probe(s)

   b. T: A-movement probe, fills Spec,TP with one DP (EPP)

```
              CP
            /    \
          C       TP
   Ā-probe(s)    /  \
               T     ...
```

A-probe with EPP,
ϕ-agreement,
nominative case, etc.

Theoretical background: C and T

Many languages exhibit an interdependence between C and T (see e.g. Fortuny, 2008 for a review), motivating a tighter connection:

- **Feature inheritance:** T features originate on C (Chomsky, 2008; Ouali, 2008; Fortuny, 2008; Legate, 2011, a.o.)
- **CT splitting:** C and T begin as a single head, with option of splitting (Martinović, 2015; Aldridge, 2015, last talk)
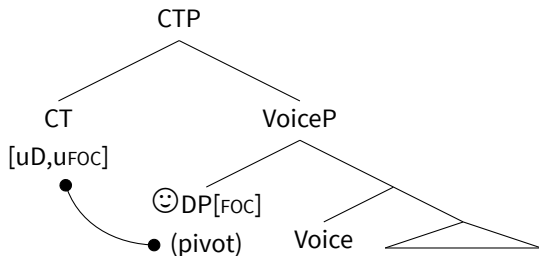
  > *"the splitting occurs in cases where a feature cannot be checked... or because there is no available position for its goal to move into."*   *Martinović (2015, p. 64)*

Proposal

(18) **Proposal:**

    a. I adopt CT splitting: **CT starts as one head**

    b. C is associated with a probe for *wh-* and *only*-phrases: [uFOC]

                                                 (cf last talk's [uWH])

    c. T is associated with a probe for a DP: [uD]   (cf last talk's [uϕ])

    d. These probes can (Case-)license their agreement targets; subsequent movement is generally optional

    e. **CT will first probe to satisfy [uD,uFOC] *together*;**
       C and T split if no [D,FOC] target is found.
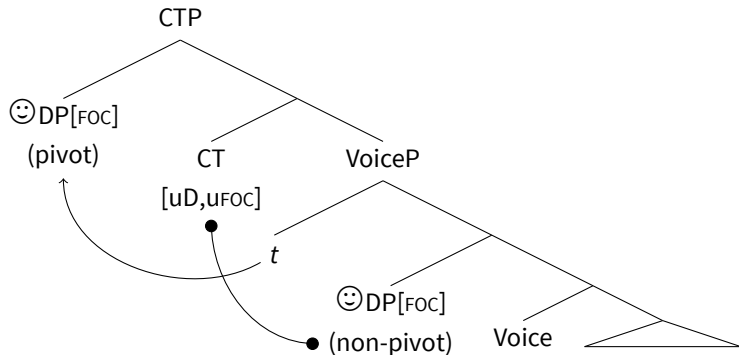   (Partially matching targets will trigger defective intervention.)

CT probes for [uD,uFOC] *together*:

```
                        CTP
                   ╱           ╲
                 CT              VoiceP
            [uD,uFOC]          ╱        ╲
                 ●      ☺DP[FOC]      ╲
                  ╲         ● (pivot)  Voice    ╱‾‾‾‾╲
                   ╲_____╱
```

Agree; license the pivot; optionally move to preverbal position

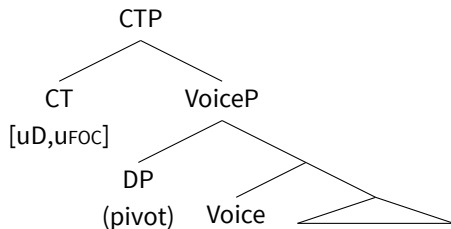# Two FOC DPs at the edge

CT probes for [uD,uFOC] together *again*:


```
                      CTP
              _____
             /                   \
      ☺DP[FOC]                     \
      (pivot)          CT          VoiceP
                    [uD,uFOC]    _____
                       •        /           \
                             t /             \
                              /               \
                        ☺DP[FOC]               \
                        • (non-pivot)   Voice   ▱
```

Agree; **license the non-pivot**; move to preverbal position

☞ Postverbal non-pivot DPs need verb-adjacency for licensing, but multiple fronting (agreeing with CT) satisfies licensing.
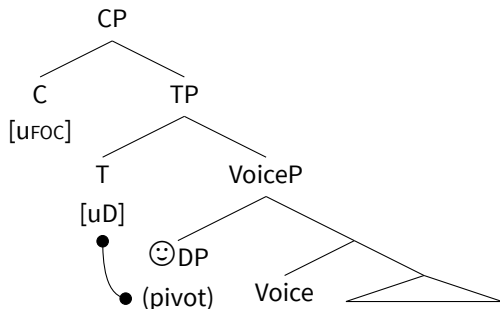
CT probes for [uD,uFOC] *together*:

☞ **If the pivot is not** [FOC]**, CT will not find any** [D,FOC] **target** at the lower phase edge, and **must split into C and T.**
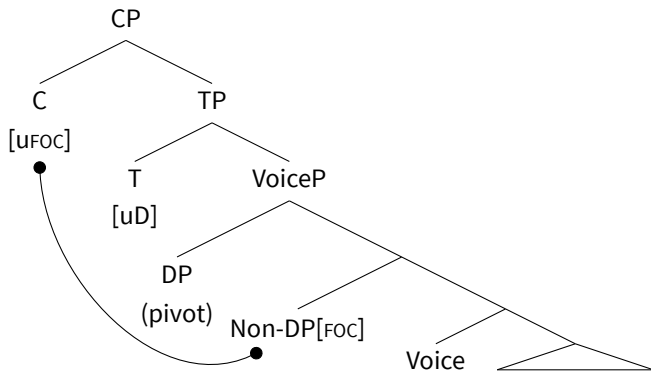
C and T splits; T probes for [uD]:

Agree; license the pivot; optionally move

C probes for [uFOC]:

Agree; move the FOC non-DP

Summary again

(19) **Summary, based on (15):**

    a.   **DP** V...   CT splits; T attracts pivot

    b.   **DP**[FOC] V...   CT attracts pivot

    c.   * DP[*wh*] **DP** V...   CT sees non-FOC pivot; CT must split; (defective intervention)

    d.   DP[*wh*] **DP**[*only*] V...   CT attracts pivot; probes again

    e.   Non-DP[*wh*] **DP** V...   CT splits; T attracts pivot; C probes

Conclusion

1. **Multiple DPs can be simultaneously extracted**, but only if both are formally focused (*wh* or *only*).

   - Motivates *initial joint probing* by [uD] and [uFOC], then separate probing;
   - ☞ In turn motivates a **CT-splitting approach** as in Martinović (2015); Aldridge (2015): **[uD] and [uFOC] must start on the same head**.

2. **The non-pivot DP *can* move**, contra Cole and Hermon (2008)

   - Takes away the primary motivation for V(oice)P-fronting;
   - Adjacency facts are better explained by a need for **licensing by adjacency** (Levin, 2015);
   - Voice tracks the pivot, which will be the **first DP attracted** (if any).

Further research

Some further directions for study:

- A- and $\overline{\text{A}}$-properties of these movements
- Multiple non-DP extractions
- Left-dislocated topics, as in Cumming (1984)

Thank you! Questions?

References I

Aldridge, Edith. 2015. Origin of the extraction restriction. LSA 2015 summer institute lecture notes.

Chomsky, Noam. 1986. *Barriers*. MIT Press.

Chomsky, Noam. 2008. On phases. In *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, ed. Robert Freidin, Carlos P. Otero, and Maria-Luisa Zubizarreta, 133–166. MIT Press.

Chung, Sandra. 2008. Indonesian clause structure from an Austronesian perspective. *Lingua* 118:1554–1582.

Clark, Robin. 1984. The syntactic nature of Logical Form: Evidence from Toba Batak. In Schachter (1984b), 9–16.

Clark, Robin. 1985. The syntactic nature of Logical Form: Evidence from Toba Batak. *Linguistic Inquiry* 16:663–669.

Cole, Peter, and Gabriella Hermon. 2008. VP raising in a VOS language. *Syntax* 11:144–197.

Cumming, Susanna. 1984. The syntax and semantics of prepredicate word order in Toba Batak. In Schachter (1984b), 17–36.

Emmorey, Karen. 1984. The intonation system of Toba Batak. In Schachter (1984b), 37–58.

Erlewine, Michael Yoshitaka, Theodore Levin, and Coppe van Urk. 2015. What makes a voice system? On the relationship between voice marking and case. In *AFLA 21: The Proceedings of the 21st Meeting of the Austronesian Formal Linguistics Association*, ed. Amber Camp, Yuko Otsuka, Claire Stabile, and Nozomi Tanaka, 51–68.

Erlewine, Michael Yoshitaka, Theodore Levin, and Coppe van Urk. to appear. Ergativity and austronesian-type voice systems. In *Oxford Handbook of Ergativity*, ed. Jessica Coon, Diane Massam, and Lisa deMena Travis. Oxford University Press. URL `http://ling.auf.net/lingbuzz/002629/current.pdf`.

Fortuny, Jordi. 2008. *The emergence of order in syntax*. John Benjamins.

References III

Guilfoyle, Eithne, Henrietta Hung, and Lisa Travis. 1992. Spec of IP and Spec of VP: Two subjects in Austronesian languages. *Natural Language & Linguistic Theory* 10:375–414.

Legate, Julie Anne. 2011. Under-inheritance. Presented at NELS 42.

Legate, Julie Anne. 2014. *Voice and* v: *Lessons from Acehnese*. MIT Press.

Levin, Theodore. 2015. Licensing without case. Doctoral Dissertation, Massachusetts Institute of Technology.

Martinović, Martina. 2015. Feature geometry and head-splitting: Evidence from the morphosyntax of the Wolof clausal periphery. Doctoral Dissertation, University of Chicago.

Ouali, Hamid. 2008. On C-to-T feature transfer: the nature of agreement and anti-agreement in Berber. In *Agreement restrictions*, ed. Roberta D'Alessandro.

Schachter, Paul. 1984a. Semantic-role-based syntax in Toba Batak. In Schachter (1984b).

Schachter, Paul, ed. 1984b. *Studies in the structure of Toba Batak*. Number 5 in UCLA Occasional Papers in Linguistics.